



CMS Distributed Processing Environment (DPE)

Greg Graham
FNAL CD/CMS

22-June-2004

<http://www.uscms.org/scpages/subsystems/gridinfo.html>



Software Under Development

- Support for Distributed/Grid Operations
 - DPE (Distributed Processing Environment)
 - Packages on top of VDT/Grid 2003
 - Specialized Monitoring
 - Security
 - Software Distribution
- Production Processing
 - CMS Production Software (RunJob)
 - Monte Carlo Processing System



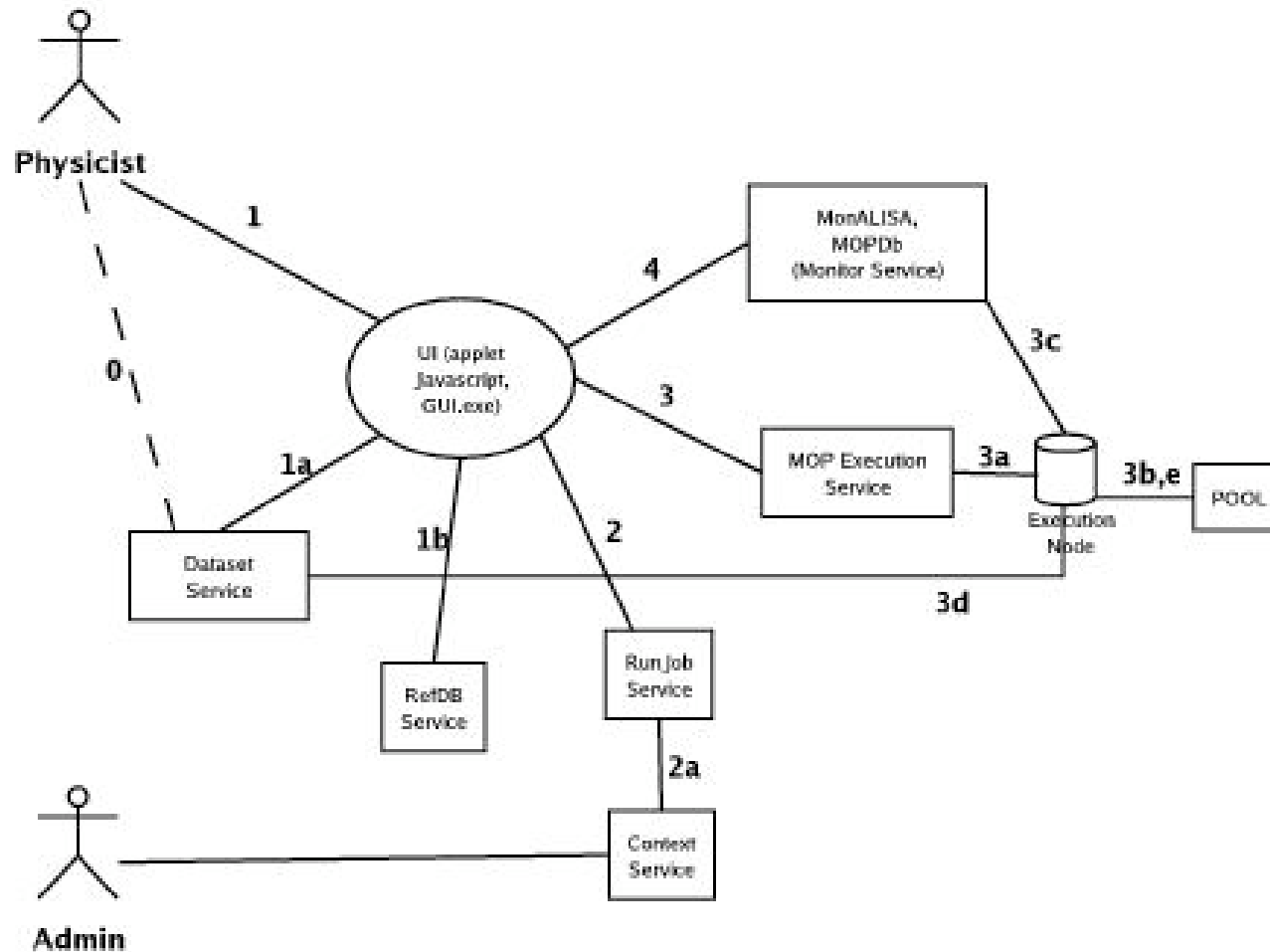
Monte Carlo Generation Service

- It is difficult for ordinary physicists currently to generate Monte Carlo or do batch style analysis in CMS without heavy support.
 - Use existing components RunJob, Clarens, and MOP to implement a system that responds well to users. Limit scope to FNAL first and then USCMS.
- Document being circulated in USCMS DPE group
 - addresses technology choices, architecture, and security.
 - DPE software assumed. (VDT/Grid2003 with CMS additions.)
 - and will be discussed further at a special DPE meeting on Thursday afternoon.



Monte Carlo Generation Service

as)



22-June-2004

Greg Graham FNAL CD/CMS



Dataset Service

- Maintains a list of datasets available in dCache at Fermilab.
 - PCP '04 production results.
 - New DSTs created on the PCP '04 data.
- Integrates several different layers of metadata
 - POOL Metadata (non-RLS version)
 - CARF Metadata
 - User/physics metadata not included
- Non-trivial preparation of the dataset
 - Continuing fallout from Objectivity architecture



Prototype Dataset Service

- Phase 1: A web page that provides dataset names, a POOL XML catalog, and application input required to use the dataset.
 - http://home.fnal.gov/~gug/cms/dataset_service.html
- Phase 2: Add some user metadata, uses RLS.
- Phase 3: Interacts well with the data handling system.
 - Yet to be designed!

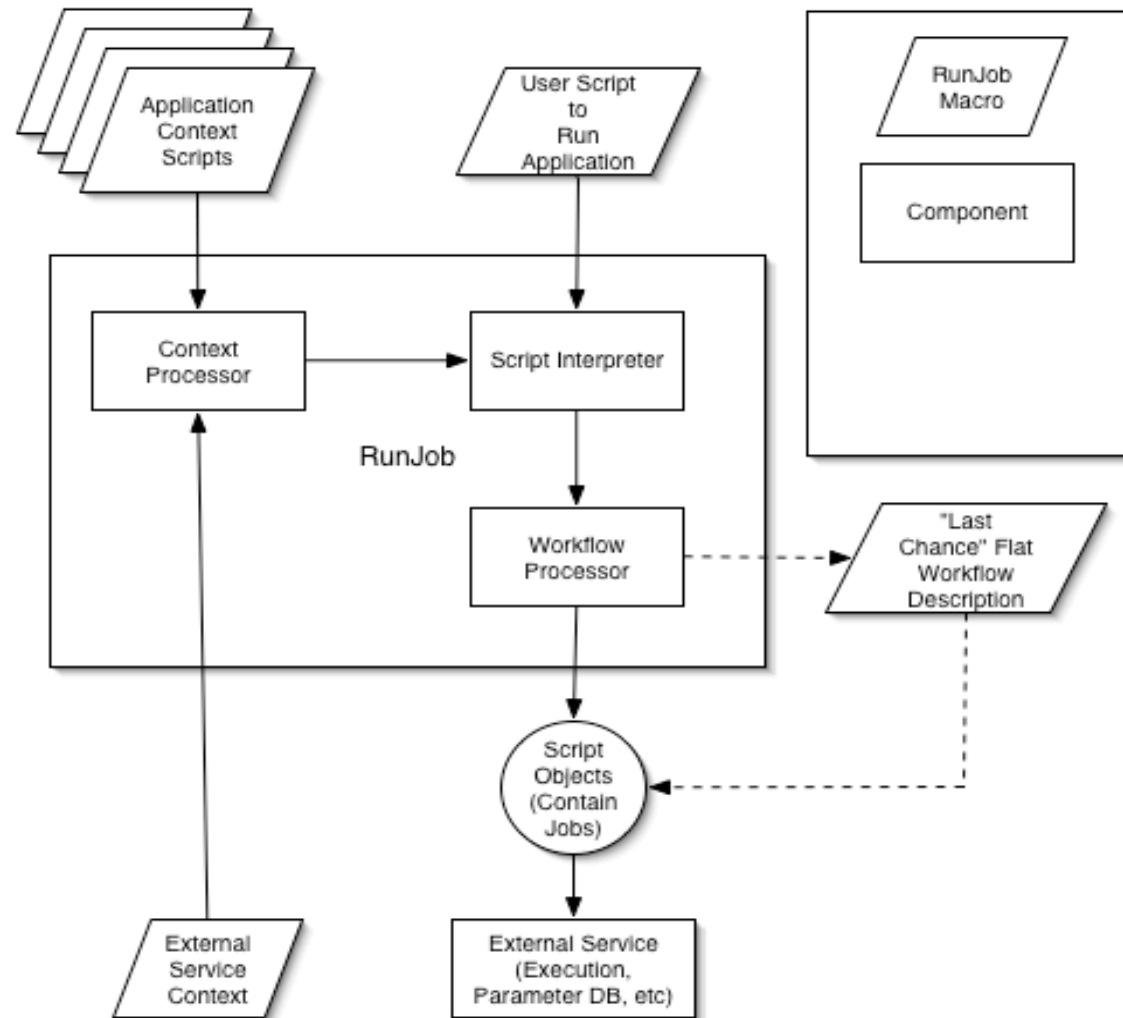


RunJob in CMS

- RunJob software is being integrated into the mainstream version of the CMS production tools.
 - CMSRunJob.py “integration” module developed, could be template for other experiments.
- RunJob software is being used for the Monte Carlo Generation Service in CMS.
 - Will be based on prototype “DC04Configurators.py” which were used successfully at UFL to make ntuple plots from DC04 DST data.



RunJob Instruction Flow



22-June-2004

Greg Graham FNAL CD/CMS

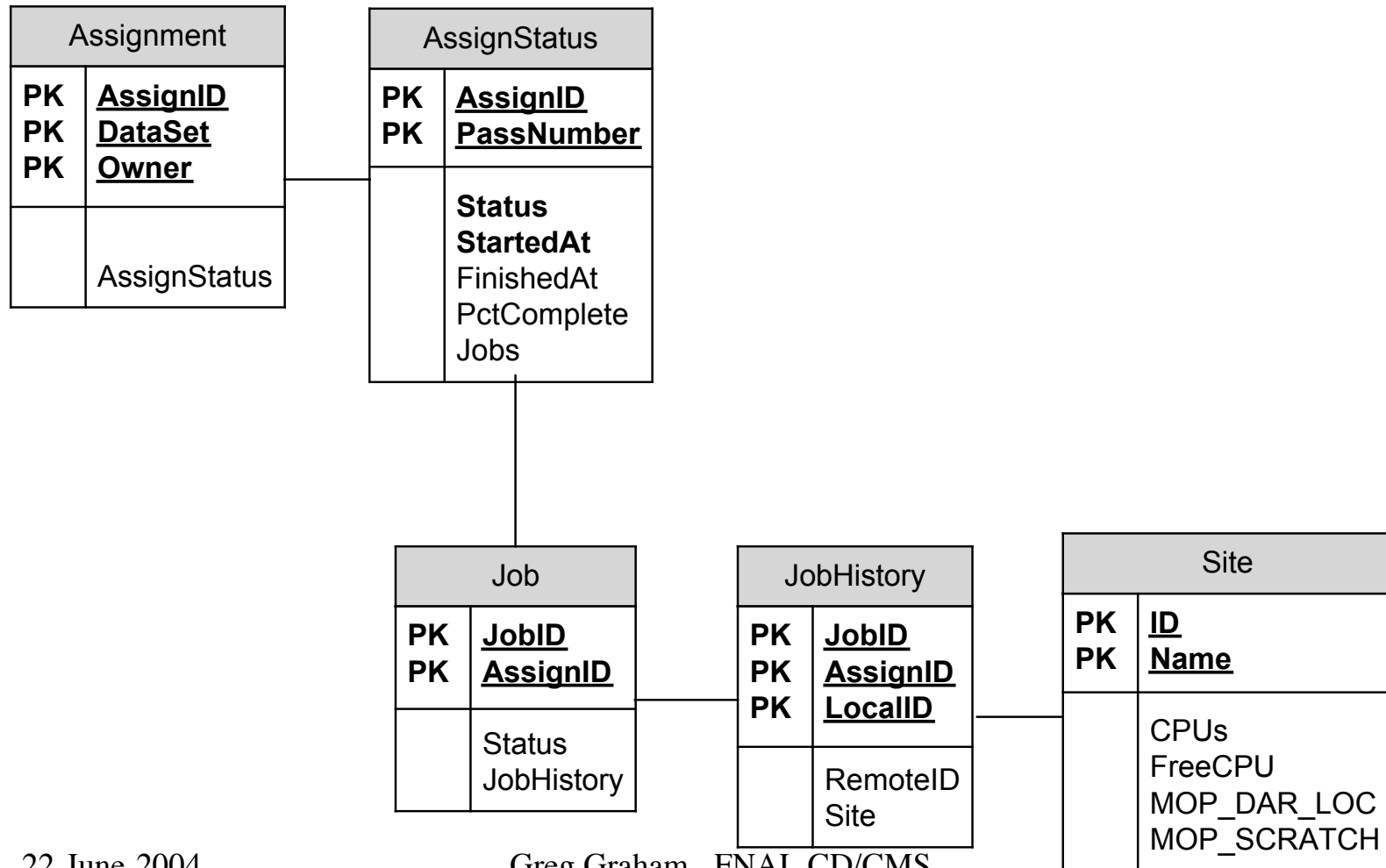


MOP and MOPDb

- All tasks done today to run production as seen by the experts: Anzar and Craig.
 - Managing Assignments
 - (status, responsibility)
 - Managing Jobs
 - (creation, submission, restart)
 - Managing Transfers
 - (locations, cksum, status)
 - Debugging
 - (when, where, how)
 - Coordination, Reporting (charts and texts)
- Production Made Easy, Click here and Produce, Your production few clicks away.. ☺



MOPDb Schema



22-June-2004

Greg Graham FNAL CD/CMS



Conclusion

- We are developing a wide range of software to help users cope with complex and distributed environments
 - Also we form strategic collaborations with organizations such as PPDG and GAE.
- We have regular DPE meetings
 - (bi-weekly on Tuesdays at 3PM)
 - minutes are kept on the project homepage.